

**MENGESAN DIF DALAM ITEM PEPERIKSAAN MATA
PELAJARAN PERDAGANGAN SIJIL PELAJARAN
MALAYSIA (SPM) SUKATAN PELAJARAN TINGKATAN
EMPAT**

ZAUBEDAH BINTI HAJI T.A. KUNCHI KOYA

**UNIVERSITI SAINS MALAYSIA
2007**

**MENGESAN DIF DALAM ITEM PEPERIKSAAN MATA
PELAJARAN PERDAGANGAN SIJIL PELAJARAN
MALAYSIA (SPM) SUKATAN PELAJARAN TINGKATAN
EMPAT**

Oleh

ZAUBEDAH BINTI HAJI T.A. KUNCHI KOYA

**Tesis yang diserahkan untuk memenuhi keperluan bagi
Ijazah Sarjana Sastera (Pendidikan)**

PENGHARGAAN

Saya bersyukur ke hadrat Allah S.W.T. kerana dengan izinNya membolehkan saya menjalankan kajian ini dengan tabah hingga selesai. Semasa menjalankan kajian, saya mendapat bantuan dan kerjasama yang tidak ternilai daripada pelbagai pihak. Saya mengambil kesempatan ini untuk merakamkan setinggi-tinggi penghargaan dan ucapan terima kasih kepada mereka berkenaan. Ucapan terima kasih terutamanya ditujukan kepada pihak Kementerian Pelajaran Malaysia, kerana telah memberikan hadiah biasiswa dan cuti sepenuh masa untuk menyiapkan tesis ini. Terima kasih juga ditujukan kepada Jabatan Pelajaran Negeri Pahang kerana membenarkan saya menjalankan kajian di beberapa buah sekolah di negeri tersebut.

Penghargaan khas saya berikan kepada penyelia saya Prof. Madya Dr Mokhtar Ismail dan Dr Wan Mohd Rani B. Abdullah selaku penyelia bersama yang telah memberikan bimbingan, panduan dan tunjuk ajar ilmiah dan professional secara berterusan sehingga membolehkan saya menyudahkan tesis ini. Tanpa dorongan, bimbingan dan tunjuk ajar yang berterusan daripada mereka adalah mustahil untuk saya menyudahkan tesis ini dengan jayanya.

Saya berterima kasih kepada pengetua-pengetua sekolah yang terlibat dalam kajian ini kerana memberikan kerjasama yang cukup baik dalam proses pengutipan data. Tanpa bantuan mereka adalah sukar bagi saya untuk mendapatkan data yang banyak bagi tujuan analisis kajian ini.

Ucapan terima kasih tidak terhingga saya tujukan kepada Encik Nordin B. Abdul Razak yang telah membantu dalam penganalisan data dan juga kepada Encik Mohd Ali B. Shamsudin kerana sering membantu dan memberi sumbangan idea dan material kepada saya dalam menyempurnakan tesis saya. Penghargaan juga ditujukan kepada kawan-kawan sama ada di USM mahupun di sekolah yang membantu dalam menyiapkan tesis ini.

Penghargaan ditujukan khas kepada suami Encik Idris Bin Kasim yang sentiasa memahami dan memberi semangat dan dorongan serta mendoakan di saat-saat akhir penyempurnaan tesis ini. Tidak ketinggalan adik-beradik, anak-anak saudara dan cucu saudara yang sentiasa mendoakan dan menjadi sumber inspirasi sepanjang menjalankan penyelidikan ini. Alhamdulillah atas sokongan semua pihak dapat saya selesaikan tesis saya dengan jayanya.

Terima kasih.

Kandungan

Muka Surat

Pengakuan
Penghargaan
Kata Pengantar
Senarai Jadual
Senarai Rajah
Senarai Singkatan
Abstrak

Bab 1

PENGENALAN

1.1.	Pendahuluan	1
1.2.	Pernyataan Masalah	4
1.3.	Soalan Kajian	9
1.4.	Hipótesis Null	10
1.5.	Rasional Kajian	11
1.6.	Tujuan Kajian	12
1.7.	Kepentingan Kajian	13
1.8.	Definisi Istilah	14

Bab 2

SOROTAN KAJIAN

2.1.	Pendahuluan	16
2.2.	Pengenalan Kepada Item Bias	16
2.3.	'Differential Item Functioning' (DIF)	21
2.4.	Pendekatan Mengesan Bias Dalam Pengukuran Dan Penilaian.	24
2.4.1.	Statistik Differential Item Functioning Mantel-Haenszel	28
2.4.2.	Teknik Delta Plot	28
2.4.3.	Teori Gerak Balas Item (IRT)	30
2.5.	Isu-isu Bias dalam Pengukuran dan Penilaian	32
2.5.1.	Bias Jantina Dalam Pengukuran dan Penilaian	33
2.5.1.1.	Keadilan jantina dalam peperiksaan	33
2.5.1.2.	Bias jantina dalam PKBS	35
2.5.1.3.	Bias jantina dalam mata pelajaran Matematik dan Sains	38
2.5.1.4.	Bias jantina dalam mata pelajaran Matematik	42
2.5.1.5.	Bias jantina dalam mata pelajaran Sains	46

2.5.1.6.	Bias jantina dalam mata pelajaran Kemahiran Hidup	49
2.5.2.	Bias Geografi dalam Pengukuran dan Penilaian	50
2.6.	Kerangka Konsep	54

Bab 3

METODOLOGI

3.1.	Pendahuluan	65
3.2.	Rekabentuk Kajian	65
3.3.	Persampelan	66
3.4.	Instrumen Kajian	67
3.5.	Kesahan dan Kebolehpercayaan Instrumen Kajian	68
3.5.1.	Kesahan isi kandungan	68
3.5.2.	Kebolehpercayaan isi kandungan	70
3.6.	Aktiviti Pungutan Data	70
3.7.	Aktiviti Penganalisisan Data	71
3.8.	Program Analisis Komputer	72
3.8.1.	Kaedah analisis data	72
3.8.2.	Kaedah mengesan bias	74
3.9.	Batasan Kajian	75

Bab 4

ANALISIS DATA DAN DAPATAN KAJIAN

4.1.	Pendahuluan	76
4.2.	Keputusan Analisis Data	76
4.2.1.	Dapatan analisis data	77
4.2.2.	Proses mengesan bias dalam item peperiksaan Perdagangan SPM	81
4.2.3.	Kaedah 'anchoring'	82
4.2.4.	Aras kesukaran item	83
4.2.5.	Kaedah Rajah 95% Tahap Keyakinan	84
4.3.	Item Berfungsi Secara Berbeza (DIF)	91
4.3.1.	Item berfungsi secara berbeza (DIF) mengikut lokasi sekolah	92
4.3.2.	Item berfungsi secara berbeza mengikut jantina	93
4.3.3.	Item pengiraan yang bertindak secara berbeza mengikut lokasi sekolah	93
4.3.4.	Item gambar rajah yang bertindak secara berbeza terhadap kelompok jantina	94
4.4.	Pengujian Hipotesis	95
4.5.	Kesimpulan	108

Bab 5

PERBINCANGAN DAN KESIMPULAN

5.1.	Pendahuluan	110
5.2.	Ringkasan Kajian	110
5.3.	Dapatan dan Perbincangan	111
5.4.	Kesimpulan	114
5.5.	Implikasi Kajian	117
5.6.	Cadangan Kajian Lanjutan	118

Bibliografi	120
-------------	-----

Lampiran

- Lampiran A Sukatan Pelajaran Mata Pelajaran Perdagangan
- Lampiran B Jadual Penentuan Ujian
- Lampiran C Ujian Pencapaian Kajian
- Lampiran D Jadual 3.1: Ringkasan Calon dan Item – Keseluruhan Responden Sebelum Diasingkan Berdasarkan Kumpulan Tertentu (Sebelum Proses Pembersihan Pespon Dilakukan).
Jadual 13.1: Item Statistic : Measure Order – Output Sebelum Proses Pembersihan Respon Dilakukan.
- Lampiran E Jadual 3.1: Ringkasan Calon dan Item – Keseluruhan Responden Sebelum Diasingkan Berdasarkan Kumpulan Tertentu (Selepas Proses Pembersihan Rspn Dilakukan)
Jadual 13.1: Item Statiastic – Measure Order – Output Selepas Pembersihan Respon Dilakukan.
- Lampiran F Peta Penentuan Dimensi Bagi 40 Item
- Lampiran G Jadual 3.1: Ringkasan Calon dan Item Bagi Kumpulan Calon Luar Bandar.
Jadual 13.1: Items Statistics: Measure Order Bagi Kumpulan Calon Luar Bandar.
- Lampiran H Jadual 3.1: Ringkasan Calon dan Item Bagi Kumpulan Calon Bandar.
Jadual 13.1: Items Statistics: Measure Order Bagi Kumpulan Calon Bandar
- Lampiran I Jadual 3.1: Ringkasan Calon dan Item Bagi Kumpulan Calon Perempuan.
Jadual 13.1: Items Statistics: Measure Order Bagi Kumpulan Calon Perempuan.
- Lampiran J Jadual 3.1: Ringkasan Calon dan Item Bagi Kumpulan Calon Lelaki.
Jadual 13.1: Items Statistics: Measure Order Bagi Kumpulan Calon Lelaki.
- Lampiran K Rajah 95% Tahap Keyakinan Menunjukkan Item-item yang Bertindak Berbeza Terhadap Kumpulan Luar Bandar dan Bandar.
- Lampiran L Rajah 95% Tahap Keyakinan Menunjukkan Item-item yang Bertindak Berbeza Terhadap Kumpulan Calon Perempuan dan Lelaki.

Senarai Jadual

Jadual		Muka surat
1	Keputusan Pepriksaan Mata Pelajaran Perdagangan SPM Daerah Temerloh tahun 2003	5
2	Perbezaan Prestasi Calon Berdasarkan Kelompok Lokasi Sekolah	85
3	Kesukaran Item bagi Kelompok Lokasi Sekolah	87
4	Perbezaan Prestasi Calon Berdasarkan Kelompok Jantina	88
5	Kesukaran Item bagi Kelompok Jantina	90
6	Item-item DIF Mengikut Kelompok Kajian	91
7	Item-item DIF Mengikut Lokasi Sekolah	92
8	Item-item DIF Mengikut Jantina	93
9	Item Pengiraan yang DIF Mengikut Lokasi Sekolah	94
10	Item Gambar Rajah yang DIF Terhadap Kelompok Jantina	94

Senarai Rajah

Rajah		Muka Surat
1	Kerangka Konsep Kajian	64
2.	95% Tahap Keyakinan bagiKelompok Lokasi Sekolah	86
3.	95% Tahap Keyakinan bagi Kelompok Jantina	89.

ABSTRAK

Kajian ini bertujuan untuk mengesan item-item ujian pelbagai pilihan yang bias atau tidak memihak kepada responden mengikut jantina dan lokasi sekolah. Responden kajian terdiri daripada pelajar-pelajar tingkatan empat daripada sekolah-sekolah menengah di negeri Pahang. Satu set ujian pencapaian mata pelajaran Perdagangan yang mengandungi empat puluh item pelbagai pilihan digunakan untuk tujuan kajian ini. Item-item tersebut dipilih daripada sukatan pelajaran Tingkatan Empat. Analisis data telah dijalankan dengan menggunakan Program BIGSTEPS. Kaedah yang digunakan untuk mengesan DIF ialah menggunakan Rajah 95% Tahap Keyakinan, di mana beberapa item yang dianggap tidak bias diambil sebagai 'anchor'. Rajah 95% Tahap Keyakinan, menunjukkan bahawa item-item yang terletak di luar garisan 'Two Standard Error' adalah dianggap sebagai bias kepada kumpulan tersebut. Hasil kajian menunjukkan terdapat item-item bias kepada kumpulan jantina dan pelajar luar bandar. Tidak terdapat item yang bias kepada pelajar bandar. Faktor-faktor yang mungkin mempengaruhi dapatan kajian ialah proses Pengajaran dan Pembelajaran, latarbelakang pelajar serta pengalaman hidup pelajar yang berbeza. Terdapat lima item yang bias kepada pelajar luar bandar. Item-item ini menjadi bias kepada pelajar luar bandar mungkin disebabkan oleh faktor tidak termasuk dalam pengalaman hidup mereka, istilah-istilah yang digunakan tidak serasi dengan mereka, item-item tersebut memerlukan pembacaan luas, dan memerlukan pengiraan serta memerlukan daya pemikiran yang tinggi. Terdapat dua item yang bias kepada pelajar lelaki. Item-item tersebut adalah item yang bergambarajah dan juga item yang bersifat pengetahuan yang memerlukan ingatan fakta, pemahaman dan pembacaan dengan meneliti maklumat yang diberikan. Terdapat lima item yang bias kepada pelajar perempuan. Kebanyakan item tersebut ialah item bergambarajah yang memerlukan keupayaan 'visualization', iaitu mereka perlu menggambarkan dalam fikiran mereka situasi yang dinyatakan dalam item tersebut bagi menyelesaikan masalah yang terdapat dalam item tersebut. Item-item yang mempunyai 'facial bias' juga tidak memihak kepada pelajar perempuan. Item yang mengandungi 'facial bias' ialah item yang lebih cenderung kepada satu-satu kumpulan yang lebih dominan sebagai contoh lelaki.

ABSTRACT

DETECTING DIF IN THE SIJIL PELAJARAN MALAYSIA (SPM) COMMERCE EXAMINATIN FROM FORM FOUR SYLLABUS

This study was conducted to detect biased items or items which are unfavorable to students in terms of their gender and school locations. Respondents consist of form four students from Pahang state secondary schools. A set of achievement test for Commerce subject consisting of forty multiple-choice items has been used for the purpose of this study. The items were from the Form Four Syllabus. Analysis of data has been done using BIGSTEPS Programme. 95% Confidential Interval diagram is used to detect DIF, where as a number of items are selected as an anchor for the diagram. 95% Confidential Interval diagram indicated that the items that are outside of the Two Standard Error line are considered bias to the group. The findings of the study indicate the existence of biased items towards gender group and rural students. However, there is no biased towards urban students. Factors which are likely to affect findings of the study are Teaching and Learning process, students' background and students' different life experiences. There are five items which are biased towards rural students. These items were bias to the students possibly because students have no prior experience pertaining to the situation in the items, terms used in the items were not compatible with the students and the items require wide reading, calculation and high thinking ability. There were two items which are biased towards male students. These items were illustrated items, knowledge-base items that need memorization fact and reading comprehension items that required thorough understanding. There were five items which are biased towards female students. Most of these items were those with illustration which require visualization ability whereby students need to visualize the situations that stated items and solve the problem that were stated in the items. Another kind of item which is not favorable to the female students is the 'facial biases. 'Facial bias is a type of bias that favors to one domain group such as male.

BAB 1

PENGENALAN

1.1. Pendahuluan

Kurikulum Baru Sekolah Menengah (KBSM) mendukung cita-cita yang murni dan unggul selaras dengan falsafah dan hasrat pendidikan kebangsaan. Penggubalan KBSM yang berhasrat untuk melahirkan pelajar-pelajar berperibadi seimbang dan harmonis dari segi intelek, emosi, rohani dan jasmani adalah berlandaskan perakuan-perakuan Laporan Jawatankuasa Kabinet Mengkaji Perlaksanaan Dasar Pelajaran (1979) dan merupakan kesinambungan daripada Kurikulum Baru Sekolah Rendah (KBSR).

Mata pelajaran Perdagangan merupakan mata pelajaran elektif bagi peringkat menengah rendah dalam rancangan KBSM selaras dengan falsafah untuk mewujudkan insan harmonis dan seimbang. Mata pelajaran Perdagangan digubal dengan mengambil kira hasrat dan aspirasi Falsafah Pendidikan Negara. (Sukatan Pelajaran Perdagangan, 1990)

Mata pelajaran ini bertujuan untuk memberi pendidikan Perdagangan dan Keusahawanan supaya pelajar mengetahui, memahami dan dapat menerapkan prinsip-prinsip dan amalan-amalan positif bagi membantunya menjadi insan yang bertanggungjawab, berpandangan jauh, rasional serta bijaksana dalam membuat sebarang keputusan, yakin diri, sanggup berdikari dan berani menghadapi persaingan dalam apa juga usaha yang diceburi berlandaskan sifat amanah dan jujur untuk kesejahteraan diri, masyarakat dan negara. Dengan mempelajari mata pelajaran ini, pelajar akan memperoleh persediaan yang mencukupi untuk melanjutkan pelajaran mereka ke

peringkat yang lebih tinggi ataupun menceburkan diri dalam pelbagai bidang kerjaya, khususnya yang berkaitan dengan perindustrian dan perniagaan (Sukatan Pelajaran Perdagangan, 1990).

Berdasarkan Sukatan Pelajaran di atas dan pengalaman sendiri penyelidik mengajar mata pelajaran ini, Perdagangan adalah dilihat sangat menarik dan dinamik. Mata pelajaran Perdagangan membantu pelajar untuk mendapatkan gred yang baik dalam Sijil Pelajaran Malaysia (SPM) agar pelajar boleh ditempatkan di IPTA atau IPTS. Di samping itu dapat juga membantu kehidupan seharian pelajar.

Apabila membuat keputusan yang penting berdasarkan markat ujian, adalah sukar untuk mengelakkan bias, yang mana mungkin akan mempengaruhi skor calon secara tidak adil. Bias wujud apabila item yang sama diberikan kepada calon yang sama kebolehan tetapi ia menghasilkan keputusan yang berbeza kepada calon daripada kumpulan yang berbeza dari segi etnik, jantina, budaya dan agama. Dalam mana-mana pengesanan bias, langkah pertama ialah, mengenalpasti sub-kumpulan yang terlibat. Kajian dan penelitian tentang bias biasanya memfokus kepada prestasi yang berbeza terhadap jantina, etnik, budaya dan kumpulan agama (Hambleton dan Rogers, 1995).

Menurut Popham (1981) jika dikaitkan dengan pengujian, bias telah menunjukkan keburukannya yang nyata. Ujian-ujian bias menghasilkan keputusan yang disalah tafsir. Apa yang menarik adalah walaupun kita telah melalui beberapa dekad dalam pengujian pendidikan di Amerika, hanya selepas sepuluh atau lima belas tahun yang lepas sahaja penguji menjadi sensitif terhadap prosedur pengujian tradisional yang mungkin mengandungi bias.

Jenis bias dalam ujian mungkin wujud dalam banyak keadaan. Kita boleh menemukan bias jantina, agama, geografi, bahasa dan lain-lain. Bias yang sedang menular dengan hebat ialah bias kaum. Ini ditambah dengan amalan pengujian di mana layanan secara individu menjadi kaku dan juga mereka menerima layanan yang tidak adil dalam masyarakat. Terdapat banyak faktor seperti ekonomi, sejarah dan sosial, yang mana individu daripada kumpulan minoriti yang mengalami masalah utama dalam pendidikan dan ini menyebabkan perkembangan remaja mereka menjadi lembab disebabkan oleh bias dalam pengujian.

Airasian (1994) pula menyatakan salah satu faktor yang mengurangkan kesahan bagi prestasi pentaksiran yang biasa ialah bias. Bias boleh wujud dalam pemerhatian dan pentaksiran secara umum. Tujuan mana-mana pentaksiran adalah untuk menentukan keputusan yang menunjukkan prestasi pembelajaran pelajar, berbanding yang lain, ciri-ciri yang tidak sesuai bagi pelajar. Apabila faktor-faktor seperti bahasa peribumi, pengalaman, jantina atau bangsa membezakan markat ujian dengan satu kumpulan pelajar yang lain, kita katakan markat ujian itu mengandungi bias. Hal ini adalah kerana penilaian yang diadakan berasaskan prestasi satu kumpulan pelajar yang dipengaruhi oleh kemasukan kriteria yang tidak relevan dan subjektif.

Menurut Popham (1999) kewujudan bias adalah perkara biasa, iaitu difokuskan bagaimana pendidik mewujudkan bias itu. Pentaksiran bias dinyatakan sebagai prosedur yang menyinggung atau penilaian secara tidak adil kepada calon disebabkan oleh ciri-ciri peribadi seperti, jantina dan etnik mereka. Pentaksiran bias dilihat sebagai pemalsuan prestasi sesetengah calon dalam ujian dan dengan itu mengurangkan kesahan prestasi calon tersebut berdasarkan markat ujian. Dua perkara utama yang menyumbang ke arah

pentaksiran bias ialah menyinggung calon dan penilaian yang tidak adil. Kita akan berpuas hati jika peperiksaan tersebut mempunyai kesan yang lain ke atas subkumpulan tertentu yang bukan semestinya bias berkaitan kandungan ujian, walaupun kesan perbezaan tersebut mungkin menjamin penelitian selanjutnya bagi kandungan peperiksaan tersebut bagi melihat jika wujudnya bias. Dalam kebanyakan hal kesan lain daripada kandungan ujian ialah calon tidak mendapat arahan secukupnya untuk menjawab peperiksaan tersebut.

1.2. Pernyataan Masalah

Pemerhatian dan penilaian sendiri yang dibuat oleh pengkaji sebagai seorang guru yang berpengalaman selama sepuluh tahun mengajar mata pelajaran Perdagangan di sekolah luar bandar mendapati keputusan pelajar bagi mata pelajaran Perdagangan tidak begitu cemerlang di peringkat SPM. Pelajar luar bandar yang majoritinya daripada kaum Melayu sangat sukar untuk memperolehi keputusan cemerlang, mereka cuma mendapat keputusan pada tahap lulus dan ada juga yang gagal mata pelajaran ini yang dikatakan mudah cuma melalui pembacaan teks dengan bantuan daripada bahan-bahan lain sebagai rangsangan. Keputusan peperiksaan ini dibandingkan dengan pelajar daripada sekolah bandar dan sekolah jenis kebangsaan. Ketidaccemerlangan mereka mungkin disebabkan mereka sememangnya tidak pandai dalam semua mata pelajaran, mereka tidak mahu belajar dalam apa jua mata pelajaran, tidak berminat dengan mata pelajaran tersebut ataupun terdapat item-item dalam peperiksaan Perdagangan SPM bias kepada mereka. Kenyataan ini disokong dengan keputusan mata pelajaran Perdagangan SPM bagi daerah Temerloh bagi tahun 2003 seperti dalam jadual di bawah.

**Jadual 1 Keputusan Peperiksaan Mata Pelajaran Perdagangan SPM Daerah
Temerloh 2003.**

BIL	SEK	KAT	A1	A2	B3	B4	C5	C6	D7	E8	%	G9	%	JUM
1.	SMK ABU BAKAR	B			1	4	7	1	26	30	73	26	27	95
2.	SMK TEMERLOH	B				4	8	3	30	29	81	17	19	91
3.	SMK DATUK BAHAMAN	LB					2	2	18	43	60	43	40	108
4.	SMK MENTAKAB	B		1	4	1	4	4	25	24	81	15	19	78
5.	SMK KUALA KURAU	LB		1	1	0	1	1	13	46	75	21	25	84
6.	SMK SERI BAHAGIA	B		1	0	0	0	0	1	12	44	18	56	32
7.	SMK SEBERANG TEMERLOH	LB	0	0	0	1	6	8	41	57	78	31	22	144
8.	SMK HWA LIAN	B		2	4	1	5	7	14	12	94	3	6	48
9.	SMK KERDAU	LB	0	0	1	0	0	1	28	31	79	16	21	77
10.	SMK SERI SEMANTAN	B	0	0	0	1	2	2	22	36	88	9	12	72
11.	SM TEKNIK	B	1	0	5	3	3	4	32	10	98	1	2	59
12.	SM(V) PUTERI	B	0	0	0	0	4	8	37	39	94	6	6	94

Kat = Kategori

LB = Luar Bandar

B = Bandar

Seperti yang diketahui item Perdagangan di peringkat SPM yang digubal untuk semua pelajar di seluruh negara adalah sama bagi semua golongan pelajar tidak mengira perbezaan agama, bangsa, tempat tinggal, latar belakang sosio ekonomi (SES), jantina dan budaya. Calon yang mengambil peperiksaan mata pelajaran ini adalah tidak sama kemampuan kerana ia dipengaruhi oleh aspek seperti latar belakang keluarga, masyarakat dan taraf hidup. Taraf sosio ekonomi dilihat dari sudut kemampuan ekonomi keluarga mereka dalam memenuhi tuntutan pendidikan semasa dan taraf sosial yang memungkinkan mereka menerima perubahan-perubahan dan pembaharuan-pembaharuan semasa sama ada dari segi pendidikan dan kehidupan seharian.

Mata pelajaran Perdagangan SPM mempunyai 22 tajuk daripada variasi topik yang pelbagai aras kesukaran (Sumber : Kementerian Pelajaran Malaysia). Terdapat tajuk-tajuk yang sukar untuk dikuasai oleh calon luar bandar seperti peruncitan besar-besaran, perniagaan luar negeri, pelaburan, keusahawanan dan bantuan-bantuan perniagaan. Tajuk-tajuk ini tidak termasuk dalam pengalaman hidup mereka. Ini dapat disokong dengan kajian yang dijalankan oleh Draba (1977) yang mana mengutarakan item-item kepada pelajar kulit hitam dan kulit putih yang berkaitan dengan 'stealing money' dan 'picnic litter'. Item-item ini memudahkan calon yang mempunyai pandangan hidup yang realistik. Memandangkan pelajar kulit hitam membesar dalam kehidupan yang realistik, item-item ini lebih memihak kepada mereka.

Pelajar perempuan sangat tidak berminat dengan topik-topik yang memerlukan pengamatan, pemerhatian dan kajian luaran, memikirkan yang lebih mendalam serta perlu untuk membuat rujukan tambahan dari sumber-sumber luar. Rujukan luaran boleh didapati dalam akhbar, majalah, TV dan internet. Topik-topik seperti pelaburan,

keusahawanan dan komunikasi memerlukan pembacaan luas, kajian luaran dan pemikiran kreatif kerana banyak perkara dan ciptaan baru wujud setiap hari. Ini membuatkan pelajar-pelajar perempuan tidak berminat terhadap item-item daripada tajuk-tajuk ini. Ini disokong oleh Kajian TIMSS-R (1998) yang menyatakan bahawa bagi item-item Matematik yang berbentuk ayat panjang dan memerlukan calon membaca dan memahami maksud (kecekapan verbal) lebih memihak kepada calon perempuan manakala item-item yang memerlukan kecekapan 'spatial' lebih memihak kepada calon lelaki. Bagi item-item ujian Sains yang bersifat abstrak dan memerlukan keupayaan 'visualization' lebih memihak kepada calon lelaki manakala item-item yang bersifat pengetahuan yang memerlukan ingatan fakta, pemahaman, pembacaan dan penelitian lebih memihak kepada calon perempuan.

Bagi calon luar bandar yang majoritinya kaum Melayu, yang mana taraf sosio ekonominya rendah agak sukar mendapatkan bahan-bahan tambahan seperti surat khabar, TV, radio, buku rujukan, majalah akademik, komputer dan alat komunikasi yang lain. Tempat tinggal mereka jauh ke pedalaman dan sukar untuk mendapat kemudahan perhubungan dan alat komunikasi yang canggih untuk membantu mereka. Mereka sukar untuk menguasai item-item daripada topik-topik perniagaan runcit besar-besaran, perniagaan luar negeri, pemasaran, bantuan-bantuan perniagaan, keusahawanan, pelaburan, konsumerisme dan etika amalan perniagaan. Ini disebabkan pengalaman hidup mereka tidak dapat membantu mereka dalam menjawab item-item tersebut. Pelajar-pelajar luar bandar tidak biasa dengan konsep-konsep yang disoal dalam item tersebut. Wright, Mead, dan Draba (1976) menyatakan sesetengah item mungkin tidak biasa kepada sesetengah calon, istilah-istilah mungkin tidak sesuai dengan kebolehan yang

diukur akan menjadi tidak serasi kepada calon dan ia merupakan bias kepada calon tersebut.

Daripada pengalaman sendiri penyelidik, mengajar dan interaksi dengan guru-guru Perdagangan di peringkat daerah, negeri dan kebangsaan kebanyakan mereka melahirkan kekecewaan terhadap jawapan pelajar yang tidak matang dalam topik-topik tertentu. Ini dilihat daripada latihan menjawab soalan peperiksaan SPM sebenar baik soalan Kertas Satu mahupun Kertas Dua. Sebagai contoh bagi topik-topik bantuan perniagaan seperti bank, promosi, insurans, komunikasi, pergudangan dan pengangkutan sentiasa mengalami inovasi yang cuma dapat diketahui daripada sumber-sumber luar daripada buku teks yang telah lama digubal. Jika item-item daripada topik-topik ini dikeluarkan dalam soalan-soalan peperiksaan SPM maka agak sukar untuk calon luar bandar menjawabnya kerana item-item tersebut tidak memihak kepada mereka.

Laporan prestasi SPM bagi mata pelajaran Perdagangan Kertas Dua bagi tahun 2002 yang dikeluarkan oleh Lembaga Peperiksaan Malaysia menyatakan secara keseluruhannya prestasi calon adalah lemah. Bilangan calon yang boleh mengingat, memahami dan mengaplikasikan konsep-konsep berkaitan kandungan mata pelajaran ini adalah rendah. Calon-calon yang mengambil mata pelajaran ini adalah calon daripada jurusan sastera (selepas 60% terbaik ditawarkan ke jurusan sains). Justeru itu, jurang ini mengakibatkan prestasi pencapaian calon agak tidak memuaskan. Calon-calon tidak dapat mengaplikasikan konsep-konsep yang dipelajari dalam mata pelajaran Perdagangan dengan kehendak soalan.

Menurut Nitko (1983) 'facial bias' juga wujud apabila item-item digubal hanya mengambil kira satu golongan yang diminati sahaja, sebagai contoh menggunakan

perkataan 'he' dalam item adalah dianggap sebagai bias kepada pelajar perempuan. Hambleton dan Rodgers (1995) menyatakan bahawa item-item mungkin menjadi bias jika terma-terma yang digunakan tidak digunakan secara meluas dalam sesebuah negara dan mungkin membawa konotasi yang berbeza dalam kawasan berbeza dalam negara tersebut.

Walaupun terdapat banyak kajian tentang bias dijalankan, tetapi kajian tentang bias dalam item mata pelajaran Perdagangan SPM belum dijalankan. Kajian ini dirasakan perlu kerana bias sebagai pembolehubah tak bersandar dapat memberi kesan kepada item-item Perdagangan SPM yang dikatakan sebagai pembolehubah bersandar. Diandaikan terdapat bias dalam item-item tersebut maka calon-calon SPM yang mendaftar mata pelajaran Perdagangan untuk SPM tidak mendapat keputusan cemerlang terutamanya bagi calon-calon luar bandar. Ini disebabkan terdapat item-item yang disoal tidak berkaitan dengan pengalaman hidup mereka dan mereka tidak biasa dengan istilah-istilah tertentu.

Adalah diharapkan penyelidikan ini dapat menjawab beberapa permasalahan yang dikemukakan. Seterusnya dapat membantu ke arah kecemerlangan keputusan mata pelajaran Perdagangan di peringkat SPM terutamanya bagi calon-calon luar bandar.

1.3. Soalan Kajian

Untuk mencapai tujuan kajian, penyelidik telah mengenalpasti soalan-soalan kajian yang perlu dijawab seperti berikut :

1. Adakah item-item dalam peperiksaan mata pelajaran Perdagangan SPM bias

kepada calon-calon luar bandar berbanding calon-calon bandar?

2. Adakah item-item peperiksaan mata pelajaran Perdagangan SPM bias kepada calon-calon lelaki berbanding calon-calon perempuan?
3. Adakah item-item pengiraan dalam peperiksaan mata pelajaran Perdagangan SPM bias kepada calon-calon luar bandar berbanding calon-calon bandar.
4. Adakah item-item bergambar rajah dalam peperiksaan mata pelajaran Perdagangan SPM bias kepada calon-calon perempuan berbanding calon-calon lelaki.

1.4. Hipotesis Null

Adalah diharapkan kajian ini dapat menunjukkan bahawa bias wujud dalam item-item mata pelajaran Perdagangan SPM, maka dengan itu hipotesis telah dibina untuk tujuan menyokong atau menolak kewujudan bias dalam item-item tersebut.

H_1 = Tidak terdapat bias dalam item ujian mata pelajaran Perdagangan SPM di antara kumpulan pelajar luar bandar dengan bandar.

H_2 = Tidak terdapat bias dalam item ujian mata pelajaran Perdagangan SPM di antara pelajar lelaki dengan pelajar perempuan.

H₃ = Tidak terdapat bias dalam item pengiraan ujian mata pelajaran Perdagangan SPM di antara pelajar luar bandar dengan bandar.

H₄ = Tidak terdapat bias dalam item bergambar rajah ujian mata pelajaran Perdagangan SPM di antara pelajar perempuan dengan pelajar lelaki.

1.5. Rasional Kajian

Walaupun pelbagai usaha dijalankan untuk mengatasi masalah kelemahan mata pelajaran Perdagangan di kalangan pelajar luar bandar oleh guru-guru, pentadbir kurikulum sekolah dan daerah namun masalah tersebut terus kekal. Guru-guru mengharapkan mata pelajaran ini dapat membantu pelajar-pelajar tersebut bagi mendapatkan gred yang baik dalam SPM. Prestasi pelajar dalam mata pelajaran ini bergantung sepenuhnya kepada cara mereka menjawab peperiksaan di peringkat nasional. Ini adalah kerana tidak terdapat penilaian yang dibuat di peringkat sekolah iaitu Penilaian Kemajuan Berasaskan Sekolah (PKBS).

Rasional kajian ini dijalankan adalah untuk menyediakan satu alternatif untuk kita mencetuskan idea-idea baru di kalangan pembina-pembina item tidak kira di peringkat mana sama ada daerah, negeri atau negara, pentadbir kurikulum pendidikan dan guru-guru dalam bidang pengukuran dan penilaian. Pembina-pembina item haruslah mengambilkira isu-isu bias dalam membina item. Selama ini isu-isu bias dalam pembinaan item tidak diambil kira ataupun tidak diketahui kewujudannya.

1.6. Tujuan Kajian

Belum pernah dibuat kajian tentang bias dalam subjek Perdagangan SPM. Tujuan kajian ini adalah untuk mengesan item bias yang terkandung dalam subjek Perdagangan SPM menggunakan Teori Gerakbalas Item (IRT) menerusi Model Rasch. Terdapat pendekatan klasik dan moden dalam mengesan item bias. Pendekatan klasik tidak digunakan kerana ia tidak mengambil kira tahap kebolehan calon berdasarkan markat item tetapi ia mengambil kira berdasarkan ujian keseluruhan. Ia menggambarkan perbezaan dalam kebolehan calon secara am sahaja.

Kajian ini bertujuan secara khususnya untuk mengesan kebolehan yang sama terhadap dua calon yang berbeza daripada kumpulan yang berbeza tempat tinggal, bangsa dan jantina. Mereka mempunyai kebarangkalian yang sama dalam memberi gerak balas yang betul kepada satu-satu item dalam peperiksaan Perdagangan SPM.

Kajian ini bertujuan untuk mengesan bias dalam item-item Perdagangan SPM. Bagi pembinaan item jika terdapat bias dalam item-item tersebut setelah diuji kepada calon maka kajian dan pembetulan perlu dilakukan. Setelah dikaji dan diperbetulkan, item-item akan diuji kepada calon sekali lagi sehinggalah unsur bias tidak terdapat dalam item-item tersebut.

Jika item yang dibina tidak mengandungi unsur bias selepas diuji dan dianalisis maka ia boleh digunakan sebagai bahan ujian. Ini setelah dilihat kepada tahap kebolehpercayaan ujian tersebut. Ia bermaksud sejauh mana ujian itu menghasilkan markat yang konsisten apabila diuji kepada calon. Tahap ini boleh dilihat melalui penganalisan data menggunakan program analisis item yang sesuai. Jika item-item

tersebut mencapai tahap kebolehpercayaannya menghampiri 1 maka kesahan item tersebut adalah tinggi dan boleh dijadikan bahan ujian dan diuji kepada calon-calon.

1.7. Kepentingan Kajian

Tujuan kajian ini adalah untuk mengesan item bias yang terkandung dalam peperiksaan mata pelajaran Perdagangan SPM terhadap calon dari luar bandar, lelaki dan perempuan. Kajian ini diharapkan dapat membantu penggubal soalan, terutamanya di peringkat nasional, agar lebih perihatin terhadap isu bias dan memberi perhatian semasa menggubal soalan. Pihak Lembaga Peperiksaan boleh mengurangkan item yang dianggap bias atau ditiadakan langsung.

Bagi guru-guru dan pembimbing harus lebih memberi penekanan kepada topik-topik yang terkandung item yang dikatakan bias. Ini kerana topik-topik tersebut terkandung dalam sukatan pelajaran dan guru-guru tidak boleh mengabaikan topik-topik tersebut. Ini dapat membantu guru merancang, menggubal dan menyusun item-item dengan lebih berkesan untuk semua pelajar dalam sub-kumpulan tersebut. Ini kerana pelajar yang mendaftar untuk subjek Perdagangan adalah daripada golongan pencapaian akademik yang sederhana dan lemah. Ini adalah kerana pelajar lemah memerlukan lebih bimbingan dan perhatian daripada guru mereka berbanding dengan pelajar pandai. Pada kebiasaannya pelajar pandai yang mendaftar mata pelajaran ini adalah atas pilihan dan minat mereka dan mereka telah bersedia dengan segala kelengkapan untuk menghadapi peperiksaan bagi menjawab item-item yang diberikan.

Bagi pelajar sederhana dan lemah yang mendaftar subjek Perdagangan untuk peperiksaan SPM dapat merasakan subjek ini tidaklah sukar jika tidak terdapat item-item

yang bias kepada mereka. Ini adalah kerana subjek ini sangat menarik dan dapat membantu mereka untuk menempatkan diri di institusi pendidikan kemahiran teknikal dan vokasional.

1.8. Definisi Istilah

Berikut ialah definisi istilah yang digunakan dalam kajian ini :

Item bias.

Sesuatu item itu bias apabila seseorang individu yang sama tahap kebolehan, daripada kumpulan yang berbeza tidak mempunyai kebarangkalian yang sama untuk menjawab item tersebut dengan betul. Ia juga dikatakan sebagai sejenis ketidaksaan yang menzalimi satu kumpulan daripada kumpulan yang lain (Angoff, 1993).

Ujian bias.

Ia dimasukkan sebagai ketidaksaan atau ralat sistematik dalam mana ujian mengukur ahli-ahli dalam kumpulan tertentu. Bias menjadi sistematik dalam erti kata ia menyebabkan ketidakadilan dalam keputusan peperiksaan bagi ahli dalam kumpulan tertentu (Camilli dan Shepard, 1994).

‘Differential Item Functioning.’

Jika setiap item ujian mempunyai fungsi gerak balas item yang sama dalam setiap kumpulan, dengan itu calon yang mempunyai kebolehan atau kemahiran yang sama akan mempunyai peluang yang sama untuk mendapat jawapan betul bagi item tersebut tanpa mengira kumpulan mereka maka item tersebut adalah tidak bias. Jika

sesuatu item itu berfungsi berbeza kepada sesuatu kumpulan dengan yang lain, ia secara terang adalah bias (Dorans dan Holland, 1993).

BAB 2

SOROTAN KAJIAN

2.1. Pendahuluan

Dalam membuat kajian tentang mengesan bias dalam item mata pelajaran Perdagangan SPM perlulah dibuat kajian tentang kajian-kajian terdahulu. Masih belum ada kajian dalam ataupun luar negara dibuat untuk mengesan item bias dalam peperiksaan mata pelajaran Perdagangan SPM. Kajian-kajian atau catatan tentang mengesan bias dalam mata pelajaran yang lain adalah berkaitan dan dapat membantu pengkaji dari segi teori dalam membuat kajian ini.

2.2. Pengenalan Kepada Item Bias

Angoff (1993) menyatakan kajian tentang item bias telah dijalankan secara serius dalam tahun-tahun 1960an. Kajian ini dijalankan untuk membina kaedah untuk mempelajari perbezaan budaya. Ia juga mengkaji prinsip-prinsip yang domain, ini adalah alasan untuk menyatakan terdapat perbezaan dalam prestasi ujian di kalangan pelajar kulit hitam dan Hispanik dengan kulit putih untuk ujian menguji kekuatan kognitif yang mana mengandungi item yang di luar domain budaya golongan minoriti. Boleh dikatakan item-item sebegini memberi peluang yang kurang kepada golongan minoriti untuk berjaya. Matlamat ujian ini yang lebih spesifik ialah item-item yang bias terhadap pelajar minoriti harus dikeluarkan daripada ujian.

Menurut Angoff (1993) perkataan bias juga menimbulkan kekeliruan dari segi semantik iaitu antara kegunaan dari segi sosial dan statistik. Untuk kegunaan statistik satu

perkataan lain harus digunakan. Akhirnya frasa ‘Differential Item Functioning’ (DIF) digunakan untuk tujuan statistik. DIF membawa maksud pemerhatian yang mudah terhadap item-item yang dikaji berlandaskan perbezaan statistik kepada kumpulan yang dicadangkan (selepas dikawal dari segi kebolehan yang berbeza).

Pada masa sekarang istilah DIF digunakan secara meluas. DIF merujuk kepada perbezaan psikometrik tentang cara bagaimana sesuatu item berfungsi untuk dua kumpulan yang berbeza. Ia juga merujuk kepada perbezaan prestasi bagi sesuatu item di antara dua kumpulan calon yang sama kebolehan. Perbandingan kumpulan yang setara adalah kritikal kerana adalah penting untuk kita membuat perbezaan di antara cara bagaimana sesuatu item mempengaruhi dua kumpulan calon. Bagi kajian ini kedua-dua istilah bias dan DIF akan digunakan mengikut kesesuaian dan kepentingan.

Bias seperti yang dimaksudkan oleh *The American Heritage Dictionary*, menyatakan sebagai “kecenderungan atau kecondongan yang menghalang penilaian yang adil”. Itulah sebabnya orang bijak pandai akan mengelakkan bias. Ia mengganggu keadilan (Popham, 1981).

Dalam perkaitan dengan pengujian, Popham (1981) menyatakan, bias adalah sesuatu yang keji. Ujian yang bias menghasilkan keputusan yang tersalah tafsir. Jenis-jenis bias adalah banyak, iaitu bias jantina, geografi, linguistik dan lain-lain. Kemungkinan bias yang paling kerap dijumpai ialah bias kaum. Terdapat banyak faktor seperti ekonomi, sejarah dan sosial yang menekan kumpulan minoriti di mana remaja mereka mengalami tragedi dalam penilaian bila prestasi mereka lembab disebabkan terdapatnya bias dalam pengujian.

Nitko (1983) menyatakan, ramai yang mempertikaikan sama ada kesesuaian sesuatu ujian itu berfokus atau bias kepada kumpulan-kumpulan tertentu. Sebenarnya maksud ujian bias adalah kabur kerana kebanyakan maksud cuma terdapat dalam kajian sahaja. Kadang-kadang pengguna ujian mempunyai lebih daripada satu jenis bias dalam pemikiran mereka, walaupun mereka tidak dapat membezakan jenis-jenis tersebut.

Mengikut Smith, Smith dan De Lisi (2001), perhatian terakhir ke atas kesahan ialah keadilan kepada semua pelajar. Isu ini berkaitan dengan bias. Sebagai guru, kita perlu berfikir yang kita sensitif kepada isu bias. Kita bekerja keras untuk memastikan pelajar yang berbeza warna kulit, etnik, jantina dan kebolehan mendapat peluang yang sama dan berjaya dalam kelas kita. Kita harus memikirkan tentang pelajar ini dalam konteks pengukuran juga. Kita harus melihat kepada pelajar-pelajar dalam 'kumpulan baru', seperti pelajar pemalu, tulisan tidak cantik, tidak suka bercakap di hadapan kelas, sukar mendapat masa untuk menyiapkan kerja rumah atau projek, tidak mempunyai komputer di rumah, tidak kreatif, tidak mempunyai ibu bapa yang menyuruh menyiapkan kerja rumah, tidak ada ibu bapa yang membantu menyudahkan projek dan tidak dapat menggunakan internet. Kita harus memfokuskan semua arahan kepada semua pelajar. Jika terdapat pelajar yang mempunyai komputer yang boleh membetulkan kesilapan ejaan dan tatabahasa berbanding dengan pelajar yang tiada komputer langsung, bagaimana harus kita membuat penilaian tentang kedua-dua pelajar ini terhadap kerja rumah berkaitan dengan penulisan, terutamanya yang memerlukan banyak bahan rujukan?

Black (1998) menyatakan, satu halangan penting dalam pengukuran sumatif dan formatif oleh guru ialah bias. Terdapat keputusan yang mengelirukan berkaitan dengan

pengukuran yang mengandung bias. Terdapat guru-guru bertindak berbeza terhadap pelajar lelaki dan perempuan, daripada kelas sosial yang berbeza atau terhadap rupa paras. Dalam hal ini guru-guru akan memberi markat yang berbeza terhadap kerja-kerja pelajar berbanding dengan yang lain. Terdapat juga bukti-bukti di mana penilaian guru dipengaruhi oleh pendapat tentang kebolehan pelajar berbanding dengan pencapaian yang sebenar.

Black (1998) berpendapat, dalam pengubalan asal ujian kecerdasan Binet, telah didapati pada satu tahap min markat pelajar perempuan adalah lebih tinggi berbanding dengan pelajar lelaki. Item-item tersebut telah dikeluarkan supaya min markat untuk kedua-dua kumpulan akan menjadi sama. Cara ini tidak dapat dipastikan sama ada ia merupakan cara yang betul atau tidak.

Terdapat banyak cara yang mana soalan mungkin menjadi tidak adil kepada pelajar-pelajar yang berbeza. Enam kemungkinan yang biasa, diterangkan di bawah (Black, 1998) :

1. Pengujian yang menggunakan soalan yang berkaitan dengan permainan mekanikal lebih cenderung kepada pelajar lelaki manakala set yang menggunakan contoh dan situasi daripada anak patung atau kerja-kerja dapur lebih cenderung kepada pelajar perempuan.
2. Soalan-soalan kurang bias apabila diambilkira tentang setereotaip jantina seperti tajuk 'impersonal' lebih cenderung kepada pelajar lelaki dan yang berkaitan dengan manusia dan implikasi personal lebih cenderung kepada pelajar perempuan.

3. Pelajar perempuan secara umumnya tidak begitu pandai dalam soalan objektif berbanding soalan esei kerana mereka tidak suka meneka (guessing) apabila tidak boleh menjawab seperti yang dilakukan oleh pelajar lelaki.
4. Pelajar perempuan tidak bernasib baik kerana mereka mengalami rasa kebimbangan semasa ujian, ini mungkin merupakan salah satu faktor yang merendahkan prestasi pelajar perempuan berbanding pelajar lelaki.
5. Soalan-soalan yang berkaitan dengan tanggapan budaya dan sosial sebagai contoh, soalan tentang orang tua yang tinggal bersendirian menjadi normal kepada budaya barat, tetapi menjadi tidak normal kepada budaya lain dan pelajar akan melihat sebagai alasan yang luar biasa kenapa orang tua disisihkan daripada keluarga.
6. Kebiasaannya, ujian akan menggunakan bahasa kelas pertengahan atau lebih cenderung kepada kanak-kanak daripada satu kelas sosial tertentu atau daripada latar belakang bahasa tertentu, berbanding dengan kumpulan yang lain.

Crocker dan Algina (1986) menyatakan, pada akhir abad ke-60 wujud isu yang menggunakan ujian untuk memilih golongan minoriti bagi mengisi kekosongan pekerjaan dan memasuki institusi pendidikan mula mendapat perhatian yang menggalakkan dalam bidang psikometrik. Walaupun kemungkinan bias dalam markat ujian adalah isu yang dipentingkan oleh pembina ujian, tetapi tiada definisi yang dirumuskan dengan jelas untuk ujian bias yang mana boleh digunakan secara objektif bagi mengesan atau menghindarkan pengaruhnya dalam situasi pemilihan. Usaha pertama dibuat ialah untuk

membina definisi ujian bias berdasarkan garis regresi untuk meramalkan prestasi markat ujian bagi calon minoriti dan kumpulan majoriti.

Berdasarkan definisi statistik, Crocker dan Algina (1986) juga menjelaskan bahawa objektif untuk ujian bias bagi pembina item ialah mereka berharap dapat membina ujian tanpa bias dalam proses pemilihan ujian tersebut. Walau bagaimanapun, penekanan ke atas perkara ini mengurang dan isu bias dalam pemilihan telah dipersoalkan dengan hebat melalui nilai-nilai sosiopolitik. Walau bagaimanapun garis regresi dibandingkan untuk kumpulan minoriti dan majoriti. Prosedur pemilihan atau peraturan boleh dirancang untuk menepati mana-mana matlamat sosiopolitik yang tertentu. Dengan ini terdapat beberapa proses dalam pemilihan dirancang yang membantu secara berbeza berkaitan dengan keadilan dalam pemilihan.

2.3. ‘Differential Item Functioning’ (DIF)

Holland dan Wainer (1993) menyatakan, bahawa analisis DIF menyediakan maklumat kuantitatif yang penting kepada kajian terhadap kesaksamaan dalam ujian atau kepada item ujian. Ia bukan sahaja maklumat yang memberitahu isu-isu tersebut, tetapi ia juga secara langsung berguna kepada persoalan mengenai perbezaan dalam prestasi kepada subkumpulan calon dan menyediakan maklumat yang berbeza daripada pengukuran kebolehpercayaan secara tradisional dan kesahan permukaan.

Istilah kesahan instrumen pengujian adalah merupakan istilah yang paling terkenal di kalangan istilah ilmu yang memudahkan penggunaan instrumen pengukuran ujian. Dalam pandangan ini ujian dicipta untuk mengukur aspek kognitif pelajar dan digunakan untuk membanding kebolehan calon-calon. Analisis DIF telah menggunakan

respon-respon yang telah diukur untuk menyatakan sesuatu tentang instrumen pengukuran itu sendiri.

Analisis DIF sepatutnya memberi gambaran secara komprehensif berasaskan maklumat tentang perbezaan prestasi calon dengan item-item ujian yang dikaji secara mendalam. Analisis DIF bagaimanapun adalah penting kepada perkara yang melibatkan kumpulan-kumpulan calon yang berlainan, contohnya kumpulan calon kulit hitam dengan kumpulan calon kulit putih atau kumpulan calon yang terdiri daripada jantina lelaki dan perempuan. Hal ini merupakan titik permulaan yang penting disebabkan wujudnya perbezaan yang ketara dari aspek sosial dalam masyarakat Amerika. Maklumat analisis keputusan ujian kurang jelas sekiranya ia tidak memberi pertimbangan yang teliti terhadap aspek pendidikan dan pengalaman calon.

Perkara terpenting dalam analisis DIF ialah ia mengesan sama ada sesuatu item itu bias atau tidak dengan tanpa dipengaruhi oleh perbezaan respon calon terhadap item yang berlainan. Bias melibatkan penilaian tentang sejauh manakah ujian mengukur apa yang sepatutnya diukur dari aspek kesaksamaan dalam ujian. DIF melibatkan kuantiti pengukuran terhadap item ujian mengikut konteks ujian secara keseluruhan. DIF bagi sesuatu ujian adalah tepat apabila tidak memberi kesan empirikial kepada calon-calon sebaliknya penekanan diberikan terhadap isu-isu pengujian lain dalam membincangkan bias dalam sesuatu ujian.

Nandakumar (1994) menyatakan definisi DIF merujuk kepada kebarangkalian yang tidak adil bagi respon yang betul bagi ujian oleh calon-calon daripada dua kumpulan apabila perbezaan kebolehan mereka dikawal. Simulasi keputusan dinyatakan sebagai langkah untuk membersihkan ujian dengan memisahkan item-item pelbagai

dimensi di bawah andaian yang mana pembina ujian berharap untuk mengkonstruksi ujian satu dimensi bagi populasi yang terpilih. Apabila dikaitkan dengan metodologi yang disediakan akan mengarah kepada pengesahan konstruk secara statistik bagi subujian yang sah yang boleh digunakan dalam pemadanan kriteria untuk analisis DIF atau bias. Metodologi ini boleh digunakan untuk padanan kriteria dalaman dan luaran.

De Ayala dan lain-lain (1999) menyatakan, DIF boleh dimaksudkan sebagai sifat item yang ditunjukkan berbeza secara statistik untuk kumpulan berbeza selepas kumpulan-kumpulan tersebut dipadankan pada kebolehan yang diukur. Sebagai contoh, data daripada dua kumpulan yang berbeza, DIF wujud apabila terdapat perbezaan dalam kebarangkalian bersyarat kepada respon betul kepada dua kumpulan yang nyata. Dicaadangkan bahawa kewujudan DIF boleh diterangkan dengan mengenalpasti data yang dikaji tidak menggambarkan populasi individu yang homogenus tetapi pencampuran data daripada pelbagai populasi atau kelas terpendam. De Ayala dan lain-lain (1999) menyatakan juga

This conceptualization of DIF hypothesizes that when one observes DIF using the current conceptualization of DIF, it is only to the degree that the manifest groups are represented in the latest classes in different proportion (p. 1).

Van de Bergh (1995) menyatakan, istilah “differential item functioning” (DIF) merujuk kepada sama ada atau tidak konstruk psikologikal yang sama dapat mengukur merentasi kumpulan yang berbeza. Jika item tidak dapat mengukur kemahiran atau subkemahiran yang sama dalam populasi yang berbeza, ia dikatakan sebagai berfungsi berbeza atau menunjukkan item itu bias. Pendekatan pelbagai tahap untuk DIF telah diperkenalkan. Dalam model tertentu pergantungan di antara pemerhatian yang berkaitan dengan kesan kelompok adalah secara terang diambil kira. Keputusan daripada model